

The Impact of Feedback to Medical Housestaff on Chart Documentation and Quality of Care in the Outpatient Setting

Donald A. Opila, MD

OBJECTIVE: To determine whether feedback from attending physicians to residents about outpatient medical records improves chart documentation and quality of care.

DESIGN: Cross-sectional study with repeated measures.

SETTING: Primary care internal medicine clinic at a metropolitan community hospital.

PATIENT/PARTICIPANTS: Fifteen interns and 20 residents.

INTERVENTION: Attending physicians reviewed at least two charts for each resident on three occasions about 4 months apart and then discussed their findings with the residents.

MEASUREMENTS AND MAIN RESULTS: Explicit criteria defined the extent of chart documentation and the comprehensiveness of care delivery. Attending physicians also made a subjective assessment of the overall quality of care. All results were converted to 0-to-1 scales. From the first to the third period, chart documentation increased from 0.60 to 0.86 ($p < .001$), but there were no significant changes in the delivery of care or in the subjective assessments of the overall quality of care.

CONCLUSIONS: Both review of residents' outpatient medical records and periodic feedback from attending physicians improve how well medical housestaff document care in the chart.

KEY WORDS: documentation; quality of care; feedback; outpatients.

J GEN INTERN MED 1997;12:352-356.

Internal medicine resident training programs are required to evaluate the competence of their trainees in the areas of medical knowledge, clinical competence, attitudes, and behaviors. Multiple methods exist for clinical competency evaluation, including the housestaff evaluation forms provided by respective boards,¹⁻⁴ the In-Training Examination (ITE),⁵ medical record review,⁶ credentialing for procedures, the Objectively Structured Clinical Exam,⁷ nurse evaluations,⁸ patient satisfaction survey in-

formation,⁹⁻¹¹ housestaff self-assessment,¹² and the Clinical Evaluation Exercise (CEX).¹³⁻¹⁴

Medical knowledge, as measured by the American Board of Internal Medicine (ABIM) Certifying Examination, correlates well with achievement on the ABIM housestaff evaluation form and ITE.¹⁵ The ABIM housestaff evaluation form may be valid for assessing overall clinical competence, but it is less useful for providing feedback in specific areas to individual residents.² The ABIM housestaff evaluation form, ITE, CEX, and faculty predictions of completeness, however, are not even moderately correlated with residents' performance of components of the physical examination.⁶ Program directors are challenged, therefore, to be innovative in finding reliable and valid methods for evaluating the clinical competence of housestaff and providing opportunities for improvement if deficiencies are identified.

Medical record review offers an attractive mechanism for evaluating clinical competence because of its ease of implementation relative to other methods such as standardized patients. The medical record documents the specific components of patient care and also demonstrates physician thought processes and outcomes of patient management. Feedback about performance to resident physicians should lead to improvement in clinical competence.

The current practice environment of managed care involves many efforts to improve quality utilizing office record review. For example, the National Council for Quality Assurance has developed the Health Employers Data Information Set (HEDIS 3.0) to ensure equitable quality-of-care standards for consumers in both public and private health care systems. Report cards developed from these data rate the quality of care in a given practice so that patients might select the best care. Individual physician performance as well as health care delivery system processes contribute to the overall report card rating.

We conducted this project to assess the impact that medical record review with periodic feedback had on housestaff completion of the medical record and on the quality of care delivered in an outpatient medical practice.

METHODS

Practice Demographics

St. Joseph's Hospital and Medical Center is a 581-bed community hospital in downtown Phoenix, Arizona. Approximately 35 housestaff (15 interns and 20 residents) in the internal medicine residency training program attend a medical clinic one half-day per week. There are approximately 1,500 patients in the primary care practice with 10,000 office visits per year.

Received from the Department of Internal Medicine, Health Services Research Group, St. Joseph's Hospital and Medical Center, Phoenix, Ariz.

Presented at the annual meeting of the Mountain West Region of the Society of General Internal Medicine, Phoenix, Ariz., February 15-17, 1995.

Address correspondence and reprint requests to Dr. Opila: Department of Internal Medicine, Health Services Research Group, St. Joseph's Hospital and Medical Center, 350 W. Thomas Rd., Phoenix, AZ 85013.

First-year residents care for an assigned panel of 25 to 30 patients and see an average of 3 to 5 patients per office session. Second- and third-year residents care for a panel of 55 to 60 patients and see 6 to 8 patients per office session. A patient is seen only by the assigned resident unless the patient schedules an urgent care appointment, goes to the emergency department, or is admitted to the hospital. Cross-coverage is provided by other medical residents or a clinic intern staffing the urgent care program, who are encouraged to communicate significant changes in medical care to the primary care residents by paging them or leaving a note in their office mailbox.

Residents are encouraged to perform a comprehensive assessment on every patient in their panel. They are given no specific guidelines on how to accomplish this requirement. Opportunities include scheduled 1-hour office visits, chart reviews independent of office visits, or using time at the end of a routine follow-up appointment. A problem list, medication sheet, and age/gender-specific health screening and maintenance sheets are part of the medical record. No system changes in chart documentation or health care delivery were implemented during the study period.

Study Design and Measurements

A cross-sectional design with repeated measures was used to study the performance of the resident physicians in this primary care practice. The focus was on measuring the performance of the practice and not the performance of individual residents. Our performance measurements included chart documentation, health care delivery, and a subjective score of overall quality as determined by the physician evaluators. Standardized criteria for chart documentation (Fig. 1) were derived from the Joint Commission on Accreditation of Healthcare Organizations accreditation manual on patient care.¹⁵ Health care delivery was implicitly assessed by the physician evaluators along multiple domains including the initial history and physical examination, medications prescribed, follow-up care, and prevention and screening (Fig. 2).

	No (score=0)	Yes (score=1)
Medical Record Organization		
1. Problem list documented		
2. Past surgery documented		
3. Primary care provider identified		
4. Medication list completed		
5. Presence / absence of allergies noted		
	< 75% of time (score=0)	≥ 75% of time (score=1)
Progress Notes		
6. Entries dated		
7. All entries signed by housestaff / provider		
8. All entries cosigned by attending		
9. Record is legible		
Physician Review Indicated		
10. Laboratory reports initialed		
11. Imaging study reports initialed		
12. Consultations / letters initialed		

FIGURE 1. Outpatient medical record documentation standards.

Baseline performance was assessed by record review in period I. At last two medical records of patients seen in the previous 3 months for each resident physician were reviewed by attending physicians. Reviewed records were scored as follows:

Chart documentation (Fig. 1). Scores were given to questions 1 through 12 based on "yes" or "no" answers. Scores were totaled and divided by 12 (number of items) to normalize the score resulting in a potential range of values from 0 and 1.

Health care delivery (Fig. 2). Questions 1 through 7 were given scores ranging from 0 to 1.0 in intervals that represented the proportion of completed chart items (0, 0.25, 0.5, 0.75, 1). Question 8 was given a score of 1 for a yes answer and a score of 0 for a no answer. Scores were totaled and divided by 8 (number of items) to normalize the score resulting in a potential range of values from 0 and 1.

Each record was then assigned a numerical score ranging from 0 (poor) to 4 (excellent) by the attending physician evaluator to give an overall subjective impression of the quality of care. The overall subjective score was normalized by dividing by 4 resulting in a potential range of values from 0 and 1.

Residents were given individualized feedback about their period I performances during a regularly scheduled, quarterly evaluation session with their faculty advisers. The faculty advisers included subspecialty physicians as well as general internists. Faculty advisers were randomly assigned to residents, and no attempt was made to ensure that the attending physician who reviewed the chart gave feedback to the resident responsible for the chart. Each resident was given a copy of the record reviews by the faculty adviser, and expectations for improvement were discussed. It was during this first evaluation session that residents became explicitly aware of the standardized review criteria.

Audit Criteria	None of time				All of time	
Score =	0	.25	.5	.75	1	
1. Is the H&P comprehensive including PMHx, SHx, FHx, Immunizations?						
2. Can you tell what medications have been prescribed for the patient?						
3. Does each progress note record patient's concerns and physical exam (if needed)?						
4. For each progress note, is a diagnostic, treatment, and follow-up plan discernible?						
5. Do progress notes reflect periodic evaluation and treatment of all diagnoses on the problem list?						
6. If there are recurrent visits for the same problem, is there documentation of improvement or reasons for deterioration?						
7. Are all preventive services implemented (i.e. PAP, Hemoccult, BSE, Mammogram)?						
	No				Yes	
Score =	0				1	
8. Is there documentation of advance directive being obtained?						

FIGURE 2. Outpatient health care delivery standards.

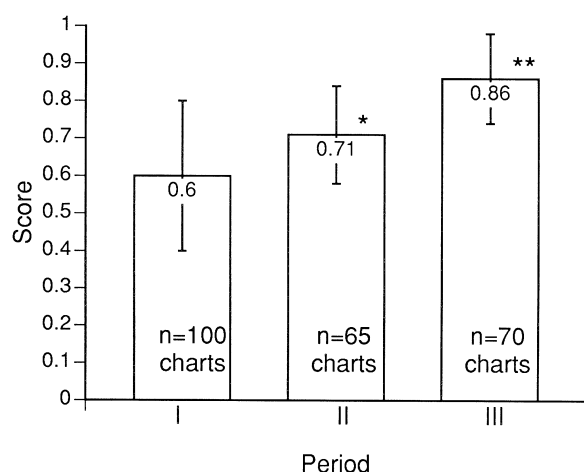


FIGURE 3. Outpatient medical record documentation results. * $p < .001$ compared with period I (baseline). ** $p < .001$ compared with period II and $p < .001$ compared with period I.

A second medical record review was performed (period II) using the same standardized criteria about 4 months after the first review. Residents were again given individualized written and verbal feedback about their chart documentation and quality of care. The second review was intended to measure the effect of feedback from period I on completion of the medical record and quality of care. A follow-up medical record review was performed in period III about 4 months after the second review to measure the effect of continued periodic feedback on chart documentation and quality of care.

Statistical Analysis

The unit of analysis for all comparisons was the medical record. We compared the medical record reviews performed in periods I, II, and III using the unpaired Student's t test when two periods were analyzed and the Tukey-Kramer test when all three periods were analyzed. We repeated the analyses for two periods using the Wilcoxon Rank-Sum Test for unpaired t tests because the proportions, particularly for the chart documentation, may not have been truly continuous variables. A repeated measures analysis of variance was done for those residents present in all three periods of evaluation and a paired Student's t test was done for those residents present in only periods II and III. Additional analyses were performed to determine how individual resident performance was affected by feedback. A p value less than .05 was considered significant. Mean values plus or minus standard deviations are reported.

RESULTS

There were 100 medical records for 36 residents in period I, 65 medical records for 32 residents in period II, and 70 medical records for 35 residents in period III. The

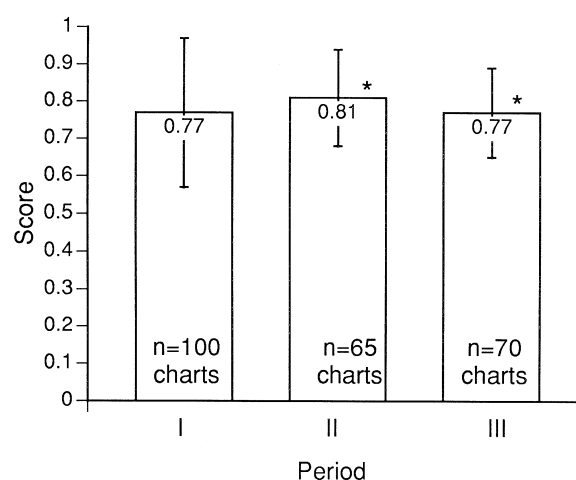


FIGURE 4. Outpatient health care delivery results. *All p values nonsignificant.

mean number of charts reviewed per resident was 4.5 (range 2–10). The medical records in period I had a mean chart-documentation score of 0.60 ± 0.20 (Fig. 3). After receiving feedback about period I performance, resident physicians improved their mean chart-documentation score during period II to 0.71 ± 0.13 ($p < .001$). A follow-up review in period III found a further increase in mean chart-documentation score (0.86 ± 0.12) compared with period I ($p < .001$) and period II ($p < .001$).

The medical records in period I had a mean health care delivery score of 0.77 ± 0.16 (Fig. 4). The overall subjective quality-of-care score, as judged by the attending physician, was 0.78 ± 0.16 (Fig. 5). A follow-up chart review in period II after residents received written and verbal feedback in period I did not find improvement in either health care delivery (0.81 ± 0.17 , $p \geq .05$) or overall subjective quality-of-care scores (0.75 ± 0.19 , $p \geq .05$) compared with period I. Follow-up chart review in period III again did not find improvement in either health care delivery scores (0.77 ± 0.16 , $p \geq .05$) or overall subjective

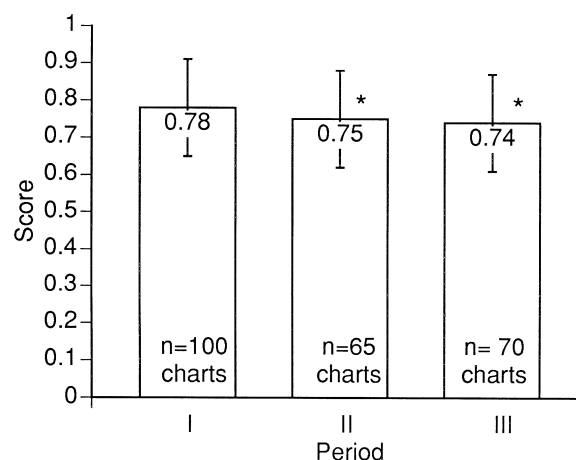


FIGURE 5. Outpatient subjective quality-of-care results. *All p values nonsignificant.

quality-of-care scores (0.74 ± 0.18 , $p \geq .05$) compared with period II. In addition, a comparison of period III with period I did not find improvement over the entire study period ($p \geq .05$). Subgroup analysis by year of training showed no differences among trainee levels. There was, however, a statistically insignificant trend toward worsening performance with progression in training.

Because the study was performed over two academic years, charts from some of the residents were not available for all three study periods. We reanalyzed the data from the eight residents who were evaluated in all three periods. The chart-documentation mean score in period I was 0.58 ± 0.21 . After receiving feedback about their period I performance, these residents increased mean chart-documentation score during period II to 0.71 ± 0.12 ($p < .001$). Follow-up review in period III found a further increase in mean chart-documentation score to 0.83 ± 0.13 compared with period I ($p < .001$) and period II ($p < .001$).

These residents' medical records in period I had a mean health care delivery score of 0.73 ± 0.15 . The overall subjective quality-of-care score was 0.75 ± 0.17 . A follow-up chart review in period II did not find improvement in either health care delivery (0.79 ± 0.14 , $p \geq .05$) or overall subjective quality-of-care scores (0.75 ± 0.20 , $p \geq .05$) compared with period I. Follow-up chart review in period III again did not find improvement in either health care delivery scores (0.75 ± 0.19 , $p \geq .05$) or overall subjective quality-of-care scores (0.75 ± 0.19 , $p \geq .05$) compared with period II. A comparison of period III quality performance with period I baseline performance did not find improvement over the entire study period ($p \geq .05$).

A similar analysis of the 26 residents who were present for both periods II and III found no differences when compared with the analysis of all the residents present for periods II and III. Further analysis of these residents by year of training revealed no differences in performance.

DISCUSSION

Various evaluation methods are used to assess the clinical competence of resident physicians. Medical record review is one such method. Clinical competence, however, has many dimensions, and auditors must be careful not to equate good medical record documentation with health outcomes. Martin and coworkers demonstrated that chart review with feedback to internal medicine residents on an inpatient, medical ward service can produce dramatic and sustained reductions (47%) in laboratory test ordering for patients.¹⁵ Although the number of tests ordered decreased significantly, no other end points were measured to determine if the quality of care was affected. Harchelroad and colleagues utilized daily record review of resident physicians during a 2-month emergency department rotation to document deficiencies in patient care and physician documentation.¹⁶ The mean percentage decrease in total errors was 10.4% when feedback was given. Sub-

group analysis to determine the relative contribution of documentation errors, such as physician signature, time seen, and disposition, and other quality-of-care errors, such as diagnosis, treatment, and medication errors, was not presented. Process measurement is easier and more reliable than audits of outcome and hence is more often the subject of investigation. It is often used as a surrogate for quality of care, which is an end point inherently difficult to measure.

The present study, done in an outpatient internal medical teaching practice, demonstrated that chart documentation improved with feedback. Health care delivery and our subjective measure of overall quality, however, did not improve. There were no "fair" or "poor" scores for health care delivery so that the lack of improvement in this dimension of quality cannot be explained by data variation alone. The health care delivery ratings attempted to measure the thoroughness of data gathering and thought processes, and whether recommended primary care measures were performed. These measurements necessarily required interpretation by the reviewer. One attending physician, who was not blinded to resident or patient identity, reviewed each medical record, which is similar to the process commonly used by managed care organizations. It is possible that there would be interrater differences if several attending physicians were to review a chart.

Abrahamson and Nyquist suggest that any assessment tool must fulfill five important conditions to effect behavior change.¹⁷ The tool must be appropriate to the object of measurement, feasible to use, economical, able to yield objective data, and used sufficiently to obtain representative data. The assessment tools in this study met these criteria with the possible exception of obtaining sufficient data to be representative.

The number of records reviewed per resident was small. Ognibene and coworkers indicate that 25 chart reviews per resident would provide a small enough measurement error to evaluate an individual resident's performance of physical examinations.⁶ Tamblyn et al. suggest that approximately 30 ratings per trainee would be required to achieve an acceptable level of reliability when using patient satisfaction ratings to assess resident performance in an ambulatory clinic.¹⁰

Generally, it is not feasible to review a large number of records per physician. Hence, feedback of aggregate performance by the entire practice is often presented to physicians along with feedback about individual performance. This is appropriate because the practice maintains a measurable quality of care even though individual physicians come and go. We adopted a similar approach in this study.

We do not know why feedback improved chart documentation but not health care delivery or our subjective measure of overall quality. Other studies have shown that improved process of care documented by explicit criteria was associated with a 5.3% reduction in 30-day mortality.¹⁸ In the current study, feedback was given simulta-

neously for chart documentation, health care delivery, and our subjective measure of overall quality, so that the mechanism of feedback itself is unlikely to be the reason for any lack of improvement. It is possible that feedback works for some behaviors and not others. There was no control group in the current study, so we may have overlooked effects from changes in residents, attending physicians, patients, call schedules, or clinic processes that impeded the ability of feedback to effect change. There were no obvious changes during the study period, however, with the exception of resident progression through the training program.

The lack of improvement in health care delivery and our subjective measure of quality could be explained if the residents did not agree that the standards were appropriate. Previous studies using feedback have failed to improve performance when there was disagreement about standards. We did not assess whether housestaff agreed with our quality-of-care standards.¹⁹⁻²¹ Grauer showed that overall quality of care delivered by family practice residents conformed to the standards of "good medical care" as judged by the subjective opinion of the author in 86.5% of cases.²² He suggests that without explicit criteria or direct resident observation, an implicit audit of process may be the best tool for monitoring resident performance and *may* be associated with a similar improvement in actual care. Using such an implicit subjective audit, we found improvement in chart documentation but failed to find improvement in health care delivery or a subjective measure of overall quality.

It may simplistic to think that feedback alone can affect physician behavior. Osman and others evaluated more than 100 trials designed to improve clinical practice and change provider behavior. They concluded that programs which combine interventions, such as preceptorships, clinical opinion leaders, patient-mediated interventions, reminders, audit, and feedback, were more likely to cause a change in clinical practice than single interventions.²³

Medical record review with periodic feedback may provide a partial solution to the problem of evaluating the clinical competence of internal medicine residents in an outpatient setting. In addition, medical record review with periodic feedback may improve how well residents document care in the patient's chart. Our results, however, failed to show that the delivery of care or a subjective measure of the overall quality of care improved when medical records were reviewed periodically and residents were given feedback about their performance. Further study is needed to identify specific indicators of quality care that are representative and measurable.

REFERENCES

1. Chapman RW. The evaluation of family practice residents: a national survey. *Fam Med*. 1993;25:650-2.
2. Haber RJ, Avins AL. Do ratings on the American Board of Internal Medicine Resident Evaluation form detect differences in clinical competence? *J Gen Intern Med*. 1994;9:140-50.
3. Thompson WG, Lipkin MJ, Gilbert DA, Guzzo RA, Roberson L. Evaluating evaluation: assessment of the American Board of Internal Medicine Resident Evaluation form. *J Gen Intern Med*. 1990;5:214-7.
4. Lancaster CY, Johnson AH, Hamadeh GN. Survey of family medicine residents evaluation methods. *Fam Med*. 1993;25:646-9.
5. Quattlebaum TG, Darden PM, Sperry JB. In-training examinations as predictors of resident clinical performance. *Pediatrics*. 1989;84:165-72.
6. Ognibene AJ, Jarjoura DG, Illera VA, Blend DA, Cugino AE, Whittier FC. Using chart reviews to assess residents' performances of components of physical examinations: a pilot study. *Acad Med*. 1994;69:583-7.
7. Day RP. Evaluation of resident performance in an outpatient internal medicine clinic using standardized patients. *J Gen Intern Med*. 1993;8:193-8.
8. Butterfield PS, Pearsol JA. Nurses in resident evaluation. A qualitative study of the participants' perspectives. *Eval Health Professions*. 1990;13:453-73.
9. Falvo D. Patient perception as a tool for evaluation and feedback in family practice resident training. *J Fam Pract*. 1980;10:471-4.
10. Tamblin R, Benaroya S, Snell L, McLeod P, Schnarch B, Abrahamowicz M. The feasibility and value of using patient satisfaction ratings to evaluate internal medicine residents. *J Gen Intern Med*. 1994;9:146-52.
11. Klessing J, Robbins AS, Wieland D, Rubenstein L. Evaluating humanistic attributes of internal medicine residents. *J Gen Intern Med*. 1989;4:514-21.
12. Dickie LD, Bass MJ. Improving problem oriented medical records through self-audit. *J Fam Pract*. 1980;10:487-90.
13. Li JT. Assessment of basic physical examination skills of internal medicine residents. *Acad Med*. 1994;69:296-9.
14. Kroboth FJ, Hanusa BH, Parker S, et al. The inter-rater reliability and internal consistency of a clinical evaluation exercise. *J Gen Intern Med*. 1992;7:174-9.
15. Martin AR, Marshall AW, Lawrence AT, Dzau V, Braunwald E. A trial of two strategies to modify the test-ordering behavior of medical residents. *N Engl J Med*. 1980;303:1330-6.
16. Harchelroad FPJ, Martin ML, Kremen RM, Murray KW. Emergency department daily record review: a quality assurance system in a teaching hospital. *Qual Rev Bull*. 1988;14:45-9.
17. Abrahamson S, Nyquist JG. Deciding how to evaluate competence. In: Lloyd JS, Langsley DG, eds. *How to Evaluate Residents*. Chicago, Ill: American Board of Medical Specialties; 1986:45-56.
18. Kahn KL, Rogers WH, Rubenstein LV, et al. Measuring quality of care with explicit process criteria before and after implementation of the DRG-based prospective payment system. *JAMA*. 1990;264:1969-73.
19. Kern DE, Harris WL, Boekeloo BO, Barker LR, Hogeland P. Use of an outpatient medical record audit to achieve educational objectives. *J Gen Intern Med*. 1990;5:218-24.
20. Lomas J, Enkin M, Anderson GM, Hannah WJ, Vayda E, Singer J. Opinion leaders vs. audit and feedback to implement practice guidelines: delivery after previous cesarean section. *JAMA*. 1991;265:2202-7.
21. Hershey CO, Goldberg HI, Cohen DI. The effect of computerized feedback coupled with a newsletter upon outpatient prescribing charges: a randomized controlled trial. *Med Care*. 1988;26:88-93.
22. Grauer K. Emergency department chart auditing in a family practice residency program. *J Fam Pract*. 1983;16:121-6.
23. Oxman AD, Thomson MA, Davis DA. No magic bullets: a systemic review of 102 trials of interventions to improve professional practice. *Can Med Assoc J*. 1995;153:1423-31.

I wish to thank the members of the Health Services Research Group (Lee Brown, MD, Philip Fracica, MD, John Heffner, MD, Robert Heiligman, MD, and Danielle Sink, MD) for providing feedback on this manuscript. I reserve special thanks for John Heffner, MD, for his statistical expertise and repeated review of my first original manuscript.